

An ac test structure for fast memory arrays

by R. C. Wong

An ac test structure (ACTS) built into fast memory arrays is required to make them truly ac-testable, with 5–10% timing accuracy. Since their ac performance is very difficult to characterize, wafer tester timing uncertainty is generally about 10–50% of a typical array access time. More accurate testers are complex and expensive; they require long development time and have complicated operation procedures. ACTS is a simpler, cheaper, and more practical means of achieving greater accuracy. An ac test structure is composed of a tunable timer and path-shifting oscillators (PSOs) built around the various access paths of the array. The timer generates the array clocks with adjustable pulse widths, and the PSOs transform time intervals into frequencies. In the future, tester accuracy will improve, but memory performance will have accelerated even more. Thus, the need for ACTS is critical and will remain so in the foreseeable future.

Introduction

Various techniques in use today partially alleviate VLSI testing problems [1–10]; for example, level-sensitive scan

design (LSSD) has reduced logic test to a combinatorial problem [1, 2]. The electronic chip-in-place technique provides a means to isolate a chip from the multiple-chip environment for purposes of diagnosis [3]. General-purpose built-in test structures have been proposed to optimize test coverage and minimize test time [4, 6, 9]. Specific test structures have also been introduced to address the test problems of embedded RAMs [5, 7, 8]. A method for generating weighted random test patterns is used to obtain complete stuck-fault coverage [10]. In all of these examples, attention has been focused on dc stuck-fault coverage and test time reduction.

Stand-alone memory array chips are generally structured, and dc testing is relatively manageable. The emerging challenge is the ac testing and timing characterization of fast memory arrays.

With the rapid improvements being made in semiconductor technology and circuit innovations, high-performance memory array chips are becoming very fast. Many timing parameters now approach the tester timing uncertainties. For example, typical array access times are ~2 ns, and typical wafer tester timing uncertainty is ~0.5 ns. Thus, ac testing and timing characterization become imprecise, with uncertainties of the order of 25%. In other words, the chip is not ac-testable, though the single-stuck-fault coverage can be optimized with some LSSD constraints. Much better ac accuracy, with errors of the order of 5–10% timing uncertainty, is needed to truly test the product performance, to characterize the design, monitor the process and defects, and manage the yield.

The tester inaccuracy is primarily due to input/output signal skews, which are very difficult to control, especially

©Copyright 1990 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

at the wafer level. The tester drivers and comparators travel long cables before reaching the chip pads. These chip pads are not directly reachable for probing, and they may be terminated on-chip with different resistors tied to various bias voltages.

Other factors that may cause timing distortions are probe contact resistance, on-chip power supply drops, on-chip temperature variations, and driver/receiver simultaneous switching. However, these factors are relatively minor, and for the most part can be taken into account.

Tester signal skew can be reduced by mounting the chip on a carrier which can be probed and by making careful and frequent calibrations of the tester. However, this is not feasible in the wafer product testing environment. Even if such engineering-mode testing operations were affordable, the accuracy provided would still be insufficient.

ACTS eliminates tester signal skews by means of a tunable timer and path-shifting ring oscillators built around the different access paths of the memory array. With this technique, tester signal skews are removed from the frequency measurements. Extremely accurate, yet simpler measurements can be made because very fine time intervals can be measured with high precision with the tunable delay elements. The objective is to reduce the timing errors to 5% or less, even in the noisy manufacturing wafer test environment. The fast memory array chips then become ac-testable, maximizing yield and minimizing ac defect escapes. Though tester accuracy will improve in the future, product performance will likely outpace tester improvement. Therefore, some built-in ac test structure is an essential element of a fast memory array.

Rationale for built-in structure

Because of the inherent timing inaccuracy of wafer testers, product ac qualification of fast memory arrays is a time-consuming process. Chips must first be diced and mounted on special chip carriers which can be probed; then their ac performance can be characterized with very elaborate manual timing measurements. The actual product performance is not known until long after the wafers are manufactured. Even when the more elaborate characterizations are completed on some selected sample of product chips mounted on probeable carriers, the timing inaccuracy is still about 10%, and the sample is relatively small. Thus, the manufacturing line process drift with respect to product performance can be only vaguely inferred from a small sample of wafers, and after a lengthy and painstaking task of performance characterization.

The ac test structure built into fast memory arrays will allow for simple and precise timing characterization on

all product chips at wafer-level test. Chip designers are always pressed to meet the design schedule, and work with constraints on chip size, power, and performance. Any test structure overhead is thus an extra burden, and, at first glance, intolerable. However, if the longer view of the total product development cycle is taken, the payback of the extra overhead at the beginning of product design is well justified. After all, the extra overhead may be no more than 10 or 20 circuits, which is a negligible number in a VLSI chip.

Without ACTS, the designer will not know whether the chip is performing to specification until much later, after the wafers are delivered. At that time, schedule pressure and the struggle to debug the design with obscure timing data will be many times more costly than the ACTS overhead.

ACTS structure

Built-in self-test structures for memory arrays have been proposed to optimize stuck-fault coverage and to reduce test time [5, 7, 8].

ACTS is designed primarily to overcome inaccuracy in the timing measurements made by the tester in the fast memory arrays. It consists of a tunable array timer and a few tunable recirculating loops traversing various access paths of the array. It must be simple enough so that it can with practicality be designed with the product; it must also be easily measurable at the wafer level. Tuning is accomplished by means of delay elements that can be controlled from external chip pads. Thus, walking strobe measurements can be performed on a chip independently of tester signal skews and inaccuracies. Also, cycle time effects on product performance can be monitored. The tunable delay facilitates the design of the array timer providing array clock pulses during regular array operations. **Figure 1** is a diagram of the ACTS structure.

• Tunable timer

Timing circuits are needed for most memory arrays. The general timing circuitry consists of some type of set/reset latch combined with a delay element to produce a single shot at the leading edge of an external clock. In general, the delay element is a chain of low-power inverters whose total delay is approximately equal to the worst-case write time. Since saturated bipolar memory cells are generally used for better immunity from the effects of soft errors or access disturbs, many inverters are required to generate the write pulse. Because of this, inverter delay does not track with the cell write time.

The tunable timer was thus developed as a feature of ACTS and as an improved substitute for the regular array timing generator. Circuits are shown in **Figure 2**. The tunable timer comprises the clock receiver latch and a tunable write-clock-delay element made of memory cell

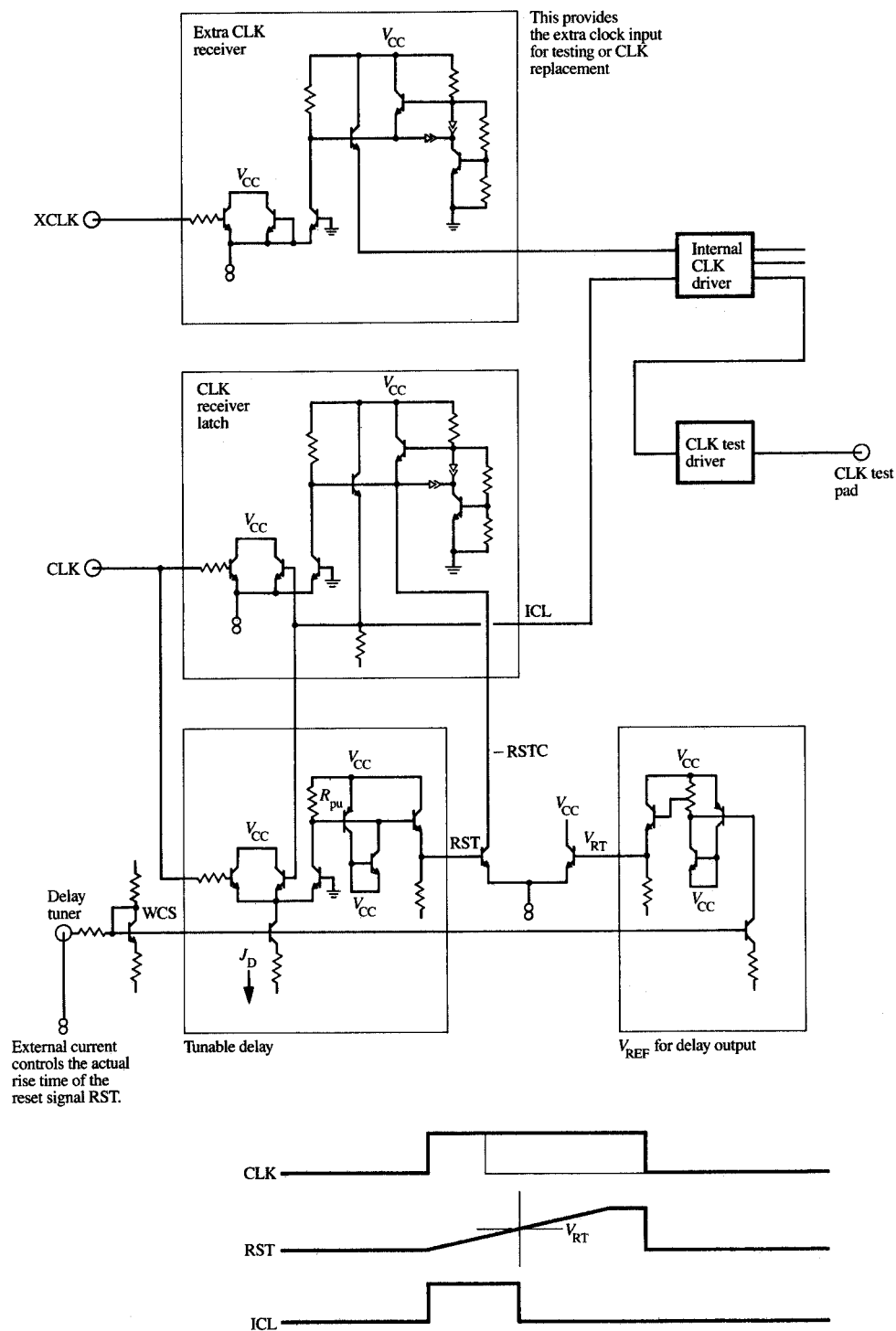


Figure 2

Tunable timer.

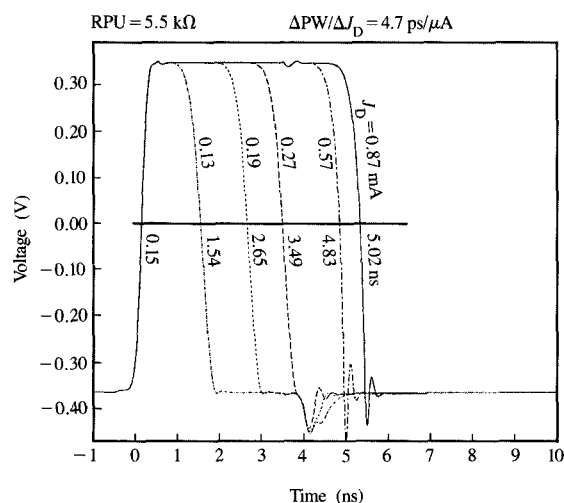


Figure 3

Tunable delay with HARPNP half cell.

Table 1 Periods of PSOs under different conditions.

-ALE	-RLE	Period
1	1	(PSOs disabled)
1	0	$P(10) = R_r + S_r + R_r + S_r$
0	1	$P(01) = R_r + A_r + S_r + R_r + A_r + S_r$
0	0	$P(00) = R_r + A_r + S_r + R_r + S_r$

R_r = rise time of reference delay

R_f = fall time of reference delay

A_r = rise time of address access path or read "1" time

A_f = fall time of address access path or read "0" time

S_r = rise time of selector

S_f = fall time of selector

$1/2 V_{be}$ below ground. Thus, the reset-dominated latch can be formed in one receiver stage, since the internal clock can feed back to the latch directly, with voltage swings compatible to those of the external clock.

The delay element is composed of the diffusion capacitance of the half-cell junction, so that the clock pulse width tracks with the cell write time.

The half cell of a CTS cell or HARPNP cell (Figure 2) is normally on, while the delay junction diffusion current J_D dictates the diffusion capacitance and, hence, the pulse width of the internal clock. The tuning pad WCS floats, and J_D is at a level which guarantees sufficient write time within the operation space of the product. The rising leading edge of the CLK sets the clock receiver latch to activate the ICL. At the same time, the current J_D is turned off, and the cathode of the half-cell junction is passively raised to V_{CC} by the pull-up resistor R_{pu} . When the cathode voltage reaches the midpoint of the full

swing, the reset signal is triggered to terminate the internal clock.

The ICL is also fed back to the delay element for two reasons:

- If the clock receiver latch is powered up to the active state, it will automatically reset to the inactive state.
- If the CLK is very short, the ICL will continue to force the reset pulse to be generated later.

Since the latch is reset-dominated, the internal clock will remain reset to the inactive level as long as the reset signal is up. Thus, if the external clock is very long, the reset signal will stay up to "chop" the long clock to the proper width.

The half-cell diffusion capacitance is used primarily because of its fine-tuning capability for ac testing. In addition to ac testability, there are other important advantages:

- Generated pulse tracks the array write time.
- One half-cell circuit can replace 10 to 20 inverters.
- Rise delay is much longer than fall delay, so that very little restore time is required for the next cycle.

Thus, if this tunable timer replaces the array clock, there is no circuit overhead. There are actually substantial circuit savings because the new clock generator is much simpler. Many stages of inverters are omitted, and there is no need of clock restore logic for short cycles. This new clock generator also provides write-time tracking and easy fine tuning, so that clock pulse changes can be accomplished without design changes.

Pulse width sensitivity with respect to the tuning current is shown in Figure 3. Since testers can control the current source much more precisely than the signal timing, much finer timing resolutions can be accomplished than with the tester walking strobos.

The extra clock XCLK provides an independent external clock to the array when the regular array clock CLK is kept inactive at the down level. Therefore, very wide or very narrow clock pulses outside the tunable range can be applied to the chip. These pulses may be needed for certain defect-screening or product-stressing procedures.

• Path-shifting oscillators

Recirculating loop frequency measurements have been widely used for timing characterizations of logic circuits. These loops were further enhanced with path-shifting oscillators (PSOs) so that the circuit rise time and fall time could be separated [11, 12]. Figure 4 shows the PSOs where the array address access time is to be acquired from the frequency measurements.

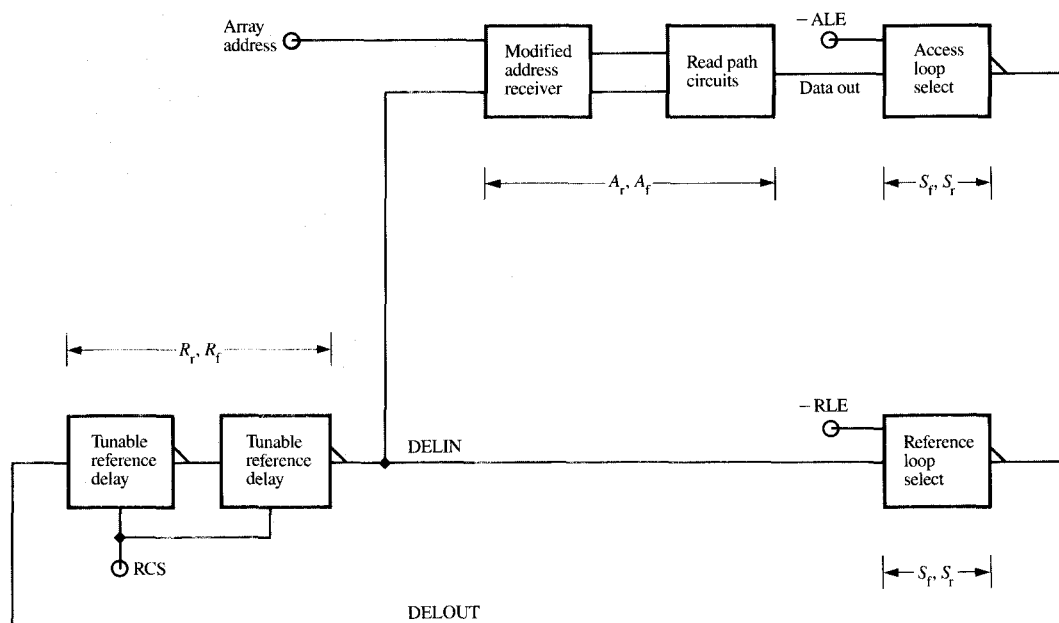


Figure 4

PSO example. Assumptions: (1) Delays of the two selectors are identical. (2) Emitter dotting does not distort delays.

The actual address input is held at the down level so that it will not interfere with the PSO recirculations. The external address loop enable ($-ALE$) signal controls oscillation through the address access path; the reference loop enable ($-RLE$) signal controls oscillation through the reference delay path. When both $-ALE$ and $-RLE$ are active at the down level, the shorter rise-time path will dominate, and the longer rise-time path in parallel is shifted off the loop due to the emitter dotting at the common output of the two loops. This is summarized in Table 1.

With respect to the PSO operations, it has been assumed that the two selectors have identical delays and that emitter dotting does not distort the actual delay. Designers must avoid circuit conditions that may invalidate these assumptions.

The address access time for reading "1" is then

$$A_r = P(00) - P(10),$$

and the access time for reading "0" is

$$A_f = P(01) - P(00).$$

These path-shifting oscillators are used in ACTS, with some other access paths multiplexed onto the common feedback bus *DELOUT* for simple and precise timing measurements.

• Tunable delay elements

In addition to the array timer delay, two other tunable delay elements are used in ACTS to allow for flexible fine-tuning of the various delays in the PSOs. The reference delay element provides the basic reference for the frequency measurements. It slows down the oscillator for fast paths so that lower-frequency-measurement equipment can be used. It also can be varied so that cycle-time effects can be observed, especially for products with tight restore-time margins.

The trigger clock delay element provides the on-chip walking strobe to measure timing parameters that cannot readily fit into a recirculating loop. The same delay element of the timer can be used for all other delay elements with some minor adjustments on the delay range. These delay elements provide very fine on-chip tuning capability that is not achievable by other means.

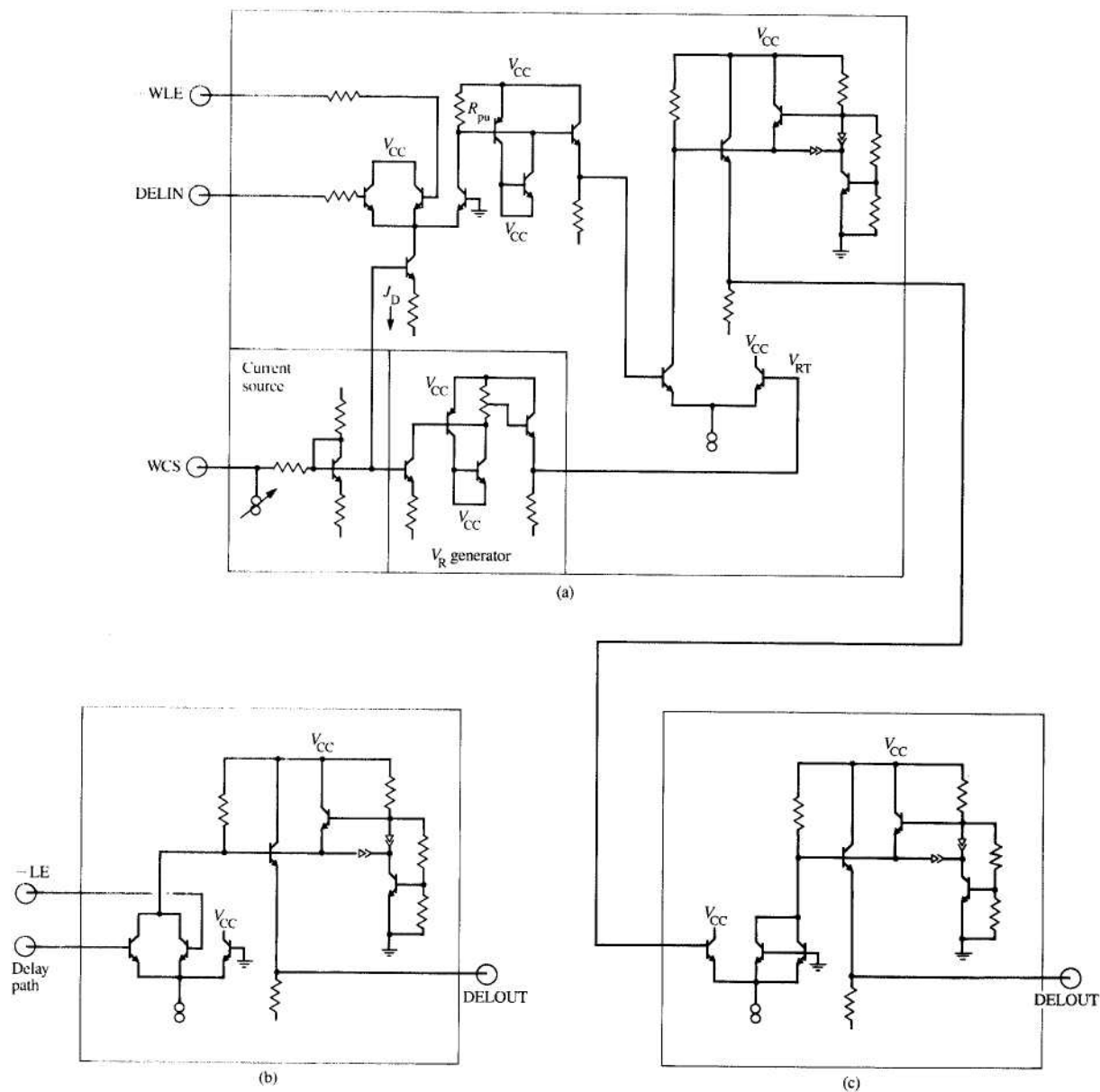


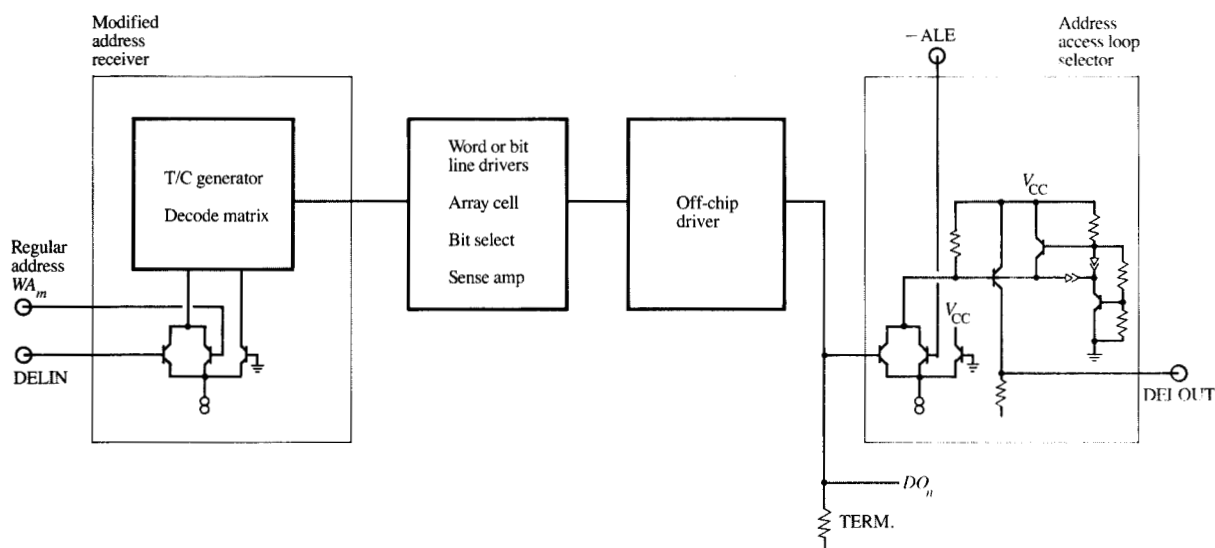
Figure 5

Array clock path circuits: (a) Tunable write clock path duplicated in the PSO. (b) Regular path selector. (c) Selector delay matching circuit.

ACTS operations

Four types of independent recirculating loops form the PSOs in the ACTS structure. In system mode, all loops are disabled. These loops are enabled in the on-chip ac test mode when the tunable delay elements are calibrated

or when a particular delay time is measured. In ac test mode, all drivers and receivers not related to the PSOs are inhibited from switching with some chip-select control signals to minimize timing distortions caused by switching noise.



- Notes: (1) Access path from WA_m to DO_n is monitored.
 (2) Cells must be written with 0 for $WA_m = 0$ and with 1 for $WA_m = 1$.
 (3) In test mode, -ALE is active down and $WA_m = 0$.
 (4) Data-out DO_n must not connect to tester probe to avoid probe loading.

Figure 6

Address access monitor.

• Write loop

The write loop (**Figure 5**) consists essentially of a duplicate of the timer delay element multiplexed onto the PSOs. Replication is a simpler and more accurate method than using extra control logic to bring the actual array timer in and out of the PSOs for frequency measurements.

This loop measures the actual array clock pulse width, minimum write pulse width, and minimum clock read time. To obtain the actual clock pulse width, the write delay tuning pad (WCS) is left floating, and the write clock pulse width (WCPW) is then derived from the period difference of the frequency measurements when both the write loop enable (-WLE) signal and the reference loop enable (-RLE) signal are active down and when only -RLE is active down:

$$WCPW = P(\text{-WLE}, \text{-RLE}) - P(+WLE, \text{-RLE}).$$

To obtain the minimum write pulse width, the array clock delay is shrunk with the control pad WCS until write operation fails. The actual pulse is then measured

with frequency measurements as before at that particular WCS current setting. The write time margin of the design is thus the difference between the regular clock pulse width and the minimum write pulse width.

In most clocked memory arrays, this same write clock controls the read timing. For example, the trailing edge of this clock may be used to latch the output drivers. Thus, by shrinking the array clock until a certain clock read access fails, the corresponding minimum read time can be derived from the subsequent PSO frequency measurements.

• Address access monitors

Since address access time is generally quoted to represent the array performance, some address access paths are configured into loops as performance monitors. The operations of the address access monitor are described in the PSO example given above, and more details are shown in **Figure 6**, where the delay path from the m th word address, WA_m , to the n th data out, DO_n , is illustrated. Only two cells are used in the loop. However, these two cells can be from many possible locations. The

other address bits not in the loop determine the location of the cells for this particular access loop. The loop cycle time is adjusted to approximately the machine cycle time by the tunable reference delay. The cells in the loop must be loaded with 0 at $WA_m = 0$, and with 1 at $WA_m = 1$. Other cells may be left at the power-up random bit pattern or may be preset to any other desired bit pattern.

Bit-pattern effects on access time can be characterized by loading the array with various bit patterns prior to the frequency measurements.

The word address receivers to be monitored are modified by adding an extra input port for the loop feedback. In regular chip operations, this feedback signal is suppressed by the feedback loop enable signals. In the on-chip ac test mode, the PSO enable signals are activated to measure access time for reading "1" or "0". For this access loop to function, there must be a feed-through path between the address receiver and the data-out driver. For example, the receiver latch and the driver latch must be flushing, and no artificial resetting may occur on the path during the recirculating operations. This delay path will include the delay of the tester probe wiring unless the probe contact with DO_n is removed during the frequency measurement, or unless a dummy driver not reaching the tester probe is used in the PSO.

- *On-chip walking strobe*

This loop provides a means of measuring other critical times that cannot fit into a ring oscillator. However, the critical time path must fit between a stimuable input and an observable output. This is illustrated in Figure 7 with scannable logic embedded in the memory array. The trigger clock (TCLK) triggers a stimulus. Simultaneously, TCLK is delayed through the tunable delay element to generate a capture clock (CCLK) at the output to capture the response. By shrinking the delay until the response fails, the minimum interval can be derived from the subsequent frequency measurements at that particular current setting of TCS. Another example is that the TCLK feeds a DATA-IN receiver latch and CCLK latches the same receiver latch. The minimum time in this case is the DATA-IN setup time.

- *Reference delay loop*

The reference delay loop provides the basic dummy delay for the path-shifting operations in the frequency measurements. The delay element comprises two tunable delay elements, which can be the half-cell diffusion capacitors, as in the other loops described above. Two inverting stages are used so that the total rise time and the total fall time are similar. This is to avoid oscillations with excessively unbalanced rise delay and fall delay.

The dummy reference delay also slows down the recirculations, so that frequency measurements can be handled more easily in the tester. The tunable feature in the reference delay path allows flexible cycle-time control, so that cycle-time effects on array performance can be observed. Cycle-time effects become critical when machine cycle time is approximate to the array access time plus the restore time, or the array clock pulse width plus the clock restore time.

Variations

ACTS has been proposed to make high-end array chips ac-testable at minimum overhead, affordable in a common VLSI chip. For a particular design, the general structure may require modifications to measure special timing parameters or to bypass special design constraints. A few potential variations are the following:

- The reference delay tuning control signal RCS may be omitted if the cycle-time effect on array performance is negligible. The reference delay can then be fixed at approximately the machine cycle time, or at the frequency range acceptable to the testers.
- The two current source controls for the write loop and for the general walking strobes, WCS and TCS, can be made common if the delay ranges are similar.
- The extra clock (XCLK) and the trigger (TCLK) can be combined if the walking strobe measurements do not interact with the regular array clock.
- XCLK can be omitted if no long clock tests are anticipated.
- Other types of performance monitors may have to be added for a particular application:
 - Compare access path for the directory array chip.
 - I/O delay monitor with just one receiver and one driver in the path.
 - Clock write through path from the clock activation to data-out valid during a write operation.

Concluding remarks

LSSD rules have been widely accepted by designers to improve the testability of dc stuck faults. However, no methodology has been established to solve the ac test problem of fast designs. Ring oscillators or path-shifting oscillators are used for the timing characterizations of stand-alone circuits in specially designed test sites. However, for fast custom product chips, such as high-end memory arrays, timings can be only vaguely inferred from those separately designed oscillators of stand-alone circuits. Thus, as semiconductor technology and circuit innovations cause chip performance to exceed tester timing accuracy, the chip becomes ac-untestable. Tester accuracy can be improved, but the process is extremely costly and is unlikely to be available in time. To simplify

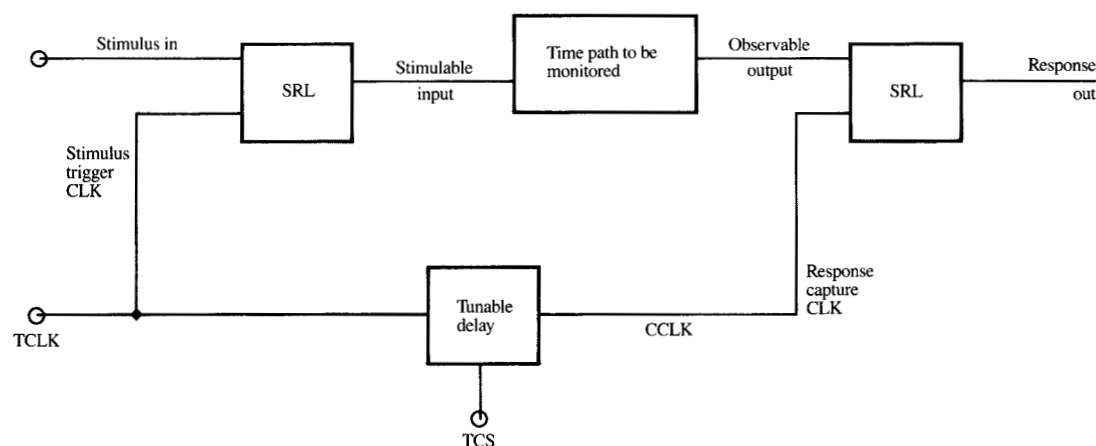


Figure 7

On-chip walking strobe.

ac verification and reduce chip development cost, an ac test structure must be built in with the chip. Better timing accuracy is provided by tunable path-shifting oscillators composed of circuits in the various performance paths of the design. This ac-testability overhead is minimal in a VLSI chip because most of the test circuits are actually part of the design.

The use of ACTS is particularly crucial at the beginning of a product cycle, when the process is new and the design objectives are very aggressive. Unfortunately, this is also the time when designers are hard pressed to meet design schedules and product objectives. Any new design overhead tends to become a low-priority requirement. Like LSSD design rules, chip-in-place test circuitry, or any other test overhead, ACTS must be included in the initial design phase so that testing will not become unmanageable in the future.

Acknowledgments

To speed up timing characterization, Product Assurance at IBM East Fishkill initiated a short-term Performance Evaluation Test Site (PETS) project. (This was a forerunner of the general-purpose ACTS, a long-term solution to the problem of tester timing inaccuracy for fast memory arrays.) The author is indebted to S. Wilson and L. Hicks for their diligent effort to correlate PETS frequency measurements with actual array product performance. The author gratefully acknowledges the

contributions of product designers P. Bunce, D. Hanson, P. Kelly, S. Koch, and G. Ritter, who helped to complete the array PETS design on a very tight schedule. The management assistance of G. Froese, F. Jones, and R. Incerto was essential to gain support from the manufacturing organization for the use of PETS to evaluate product wafers.

References

1. E. B. Eichelberger and T. W. Williams, "A Logic Design Structure for LSI Testability," *Proceedings of the 14th Design Automation Conference*, 1977, pp. 462-468.
2. S. DasGupta, R. G. Walter, and T. W. Williams, "An Enhancement of LSSD Structure and Its Applications to Non-LSSD Logic," *Proceedings of the IEEE 11th International Symposium on Fault-Tolerant Computing*, June 1981, pp. 32-34.
3. P. Goel and M. T. McMahon, "Electronic Chip-in-Place Test," *Proceedings of the 1982 IEEE International Test Conference*, pp. 83-90.
4. M. H. McLeod, "Test Circuit for Delay Measurements on an LSI Chip," U.S. Patent 4,392,105, July 1983.
5. M. H. McLeod, "Test Circuit for Turn On and Turn Off Delay Measurements," U.S. Patent 4,489,272, December 1984.
6. L. T. Wang and E. J. McCluskey, "Concurrent Built-In Logic Block Observer (CBILBO)," *Proceedings of the 1986 International Symposium on Circuits and Systems*, pp. 1054-1057.
7. S. K. Jain and C. E. Stroud, "Built In Self Testing of Embedded Memories," *IEEE Design & Test of Computers* 3, 27-37 (October 1986).
8. L. T. Wang and E. J. McCluskey, "Built-In Self-Test for Sequential Machines," *Proceedings of the 1987 IEEE International Test Conference*, pp. 334-341.

9. P. H. Bardell and W. H. McAnney, "Built-In Test for RAMs," *IEEE Design & Test of Computers* **5**, 29-36 (August 1988).
10. Y. Nishimura, M. Hamada, H. Hidaka, H. Ozaki, and K. Fujishima, "A Redundancy Test Time Reduction Technique in 1 Mbit DRAM with a Multibit Test Mode," *IEEE J. Solid State Circuits* **24**, 43-49 (February 1989).
11. M. Nicolaidis, "Self-Exercising Checkers for Unified Built-In Self-Test (UBIST)," *IEEE Trans. Computer Aided Design* **8**, 203-218 (March 1989).
12. J. A. Waicukauski, E. Lindbloom, E. B. Eichelberger, and O. P. Forlenza, "A Method for Generating Weighted Random Test Patterns," *IBM J. Res. Develop.* **33**, 149-161 (March 1989).

Received May 29, 1989; accepted for publication September 12, 1989

Robert C. Wong *IBM General Technology Division, East Fishkill facility, Route 52, Hopewell Junction, New York 12533.* Dr. Wong, who joined IBM in 1966, is a senior engineer currently working on the development of advanced cache memory. He earned his diploma in physics in 1963 from Chung Chi College, Hong Kong, and his M.S. degree in solid state science (1966) and Ph.D. degree in physics (1971), both from Syracuse University. He has worked in areas of semiconductor process development, neural network analysis, functional memory array design, microprocessor design, design rules automation, and advanced memory array development. Dr. Wong, who has achieved his eighth IBM invention plateau, received an IBM Outstanding Technical Achievement Award in 1986 for the introduction and development of the PDEX program in the EDS design system. He was a visiting lecturer at the Chinese University of Hong Kong in the academic year of 1981-1982, when he was on sabbatical leave from IBM.